Haploid Selection in “Diploid” Organisms

Simone Immler

School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom; email: s.immler@uea.ac.uk

Abstract

Evolutionary rates and strength of selection differ markedly between haploid and diploid genomes. Any genes expressed in a haploid state will be directly exposed to selection, whereas alleles in a diploid state may be partially or fully masked by a homologous allele. This difference may shape key evolutionary processes, including rates of adaptation and inbreeding depression, but also the evolution of sex chromosomes, heterochiasmy, and stable sex ratio biases. All diploid organisms carry haploid genomes, most notably the haploid genomes in gametes produced by every sexually reproducing eukaryote. Furthermore, haploid expression occurs in genes with monoallelic expression, in sex chromosomes, and in organelles, such as mitochondria and plastids. A comparison of evolutionary rates among these haploid genomes reveals striking parallels. Evidence suggests that haploid selection has the potential to shape evolution in predominantly diploid organisms, and taking advantage of the rapidly developing technologies, we are now in the position to quantify the importance of such selection on haploid genomes.

Keywords

ploidy, gametic selection, biphasic life cycle, sex chromosomes, imprinting, mitochondrial genome
INTRODUCTION

Ploidy is defined as the number of homologous chromosomes present in a cell and varies across taxa. The most commonly encountered states among organisms are haploidy (one copy of chromosomes) and diploidy (two copies of homologous chromosomes). The terms haploid and diploid were first coined in 1905 by the German botanist Eduard Strasburger in a publication describing the processes of meiotic divisions in several plant species (Strasburger 1905). Strasburger (1894) was also the first to recognize that the two alternating phases were an inevitable consequence of sexual reproduction in eukaryotes. The necessary alternation of diploid and haploid phases during sexual reproduction (Kirkpatrick 1994, Mable & Otto 1998) allows for selection to act during both phases. The rates of evolution differ strikingly between a diploid and a haploid genome (see the section titled Haploid Selection in Theory), and taking both phases into account when studying diploid organisms will improve our understanding of evolution.

In addition to the haploid genomes found in gametes, predominantly “diploid” organisms contain a range of other haploid genomes (Figure 1), and the study of these genomes may yield more accurate predictions for evolutionary dynamics. The diversity of “haploid” genomes includes the genomes rendered functionally haploid through imprinting and targeted silencing of one allele, the haploid genomes of unmatched sex chromosomes in organisms with heterogametic sex determination, and the haploid genomes of organelles such as mitochondria and plastids. The genetics of some of these haploid genomes are well studied, whereas others are still a conundrum. It may be worth zooming in on these understudied genomes to fully understand the role of haploid genomes in the evolution of “diploid” organisms.

Although the literature includes a substantial body of theoretical work on the relative importance of selection occurring in haploid and diploid genomes, particularly in biphasic life cycles, the empirical evidence is scarce. Several reasons may explain the current lack of data for the impact of a biphasic life cycle on evolutionary processes. For one, much of our focus is directed toward diplontic organisms, which spend most of their life cycle as diploids and exhibit only a very short haploid gametic phase. Furthermore, technical limitations may have contributed to the current scarcity of empirical evidence for the impact of haploid selection in diplontic organisms. This may...
be the right time to revisit the topic and improve our understanding by taking novel approaches made possible with the advent of ever-improving sequencing and other molecular technologies (see also the section titled Genomic Signatures of Haploid Gametic Selection) and by expanding our investigations across a wider range of taxonomic groups.

HAPLOID SELECTION IN THEORY

Selection acting on haploid and diploid genomes in the same organism changes the evolutionary dynamics in a number of ways. First of all, alleles expressed under haploidy will be directly exposed to selection, whereas alleles expressed in a diploid genome may be partially or completely masked by dominance and hence escape selection (Crow & Kimura 1970). In addition, the mutation rate in a diploid genome is generally assumed to be higher than in a haploid genome because of the higher number of copies present in a population. These differences between haploid and diploid genomes in turn may cause variation in expression and result in genomic conflicts, particularly if these differences in ploidy coincide with differences in the expression context such as different tissues, life stages, and/or sexes. The potential evolutionary consequences resulting from the coexistence of haploid and diploid genomes have been extensively assessed in theoretical studies; an overview of these studies is provided here.

The masking effect in diploid genomes is a key difference from haploid genomes and may affect evolution and the underlying change in allele frequencies in two fundamental aspects: (a) Positive selection favoring beneficial alleles may be hampered in a diploid genome if these are not completely dominant (Orr & Otto 1994), and (b) purifying selection will be less efficient in removing recessive deleterious alleles in a diploid genome (Crow & Kimura 1970). The difference in the efficiency of positive selection due to ploidy may directly affect the rate of adaptation and often accelerates the spread of beneficial de novo mutations in haploid populations compared with diploid populations (Orr & Otto 1994). Similarly, the effective removal of deleterious mutations has potentially important implications for the genetic load in a population (Charlesworth & Charlesworth 1992), which is of particular interest in the case of inbreeding depression (Charlesworth & Charlesworth 1987, Losdat et al. 2014). In fact, even a short window of haploid selection may effectively remove deleterious mutations from a population and thereby reduce the genetic load (Otto et al. 2015). The efficiency of haploid selection described here has far-reaching consequences and may affect the evolution of diplontic organisms in many ways.

The assumption of an increased mutation rate in diploid genomes is largely based on the fact that a diploid genome contains double the number of nucleotides compared with a haploid genome. In fact, the doubled mutation rate in diploid genomes is an underlying assumption of many theoretical models investigating the evolution of biphasic life cycles and ploidy (e.g., Crow & Kimura 1965, Kondrashov & Crow 1991). However, a recent comparison of the rates and types of mutations occurring in haploid and diploid yeast strains showed that the difference in mutation rates between the two is more complex than previously assumed (Sharp et al. 2018). In fact, haploid strains seemed more susceptible to single-nucleotide mutations, suggesting that the assumption about mutation rates in diploid and haploid genomes may need careful revisiting.

Variation in ploidy of genomic regions, chromosomes, or even entire genomes is often discussed in the context of genomic conflict, particularly if alleles are exposed to differential or antagonistic selection. Sexually antagonistic selection and the potential for intralocus sexual conflict may lead to targeted silencing through imprinting of one allele in a specific sex (e.g., Day & Bonduriansky 2004, Arnvist & Rowe 2005), rendering this locus functionally haploid in this sex. Similarly, genetic sex determination systems often involve ploidy differences between males and females either for sex chromosomes or the entire genome (Bachtrog et al. 2014). Such ploidy differences between the sexes can be driven by sexually antagonistic selection (Immler & Otto...
However, whether sex differences in ploidy are causing such conflicts or whether they may help mitigate existing conflicts is not entirely clear, and the answer may vary for the different scenarios.

Biphasic life cycles with alternating diploid and haploid phases offer an additional opportunity for genomic conflict between the phases even in predominantly diploid organisms (Immler & Otto 2018). One of the first to specifically consider selection at the haploid gametic phase was Haldane (1924), who showed that selection at this level is particularly effective compared with selection at the diploid stage. Haldane’s model was the basis for deterministic and stochastic models showing that alternating phases of haploid and diploid selection (Scudo 1967, Hartl 1977) and antagonistic selection across the ploidy levels and between the sexes (Ewing 1977, Immler et al. 2012) can lead to the maintenance of stable genetic polymorphisms. The potential conflict between the diploid gamete-producing organism and its haploid gametes arising from such antagonistic selection has specifically been discussed in the context of sperm competition. A general prediction is that such a conflict should lead to the silencing of the haploid gametic genome to reduce the risk of a conflict (Haig & Bergstrom 1995). The reason for this prediction is a potential conflict of interest over controlling the sperm phenotype and hence the ability to successfully outcompete rival sperm. Two theoretical studies, one assuming diploid male control (Parker 1993) and one assuming haploid gametic control (Parker & Begon 1993) over the evolution of sperm traits directly related to their competitive ability, showed that the potential conflict between the diploid male and its haploid sperm over ejaculate expenditure strongly affects the evolutionary stable strategies of male reproductive investment. In contrast, a more recent theoretical study predicted that the fierce competition among sperm within the ejaculate of a male in fact may increase the rate of haploid-expressed genes due to an increased fixation rate of advantageous alleles (Ezawa & Innan 2013). Furthermore, the benefits of purifying selection may favor the evolution of haploid selection, unless alleles are under poloidally antagonistic selection, in which case females will evolve to reduce haploid selection (Otto et al. 2015).

The role of haploid selection in combination with sexually antagonistic selection may affect a range of evolutionary processes. Sex-specific haploid selection in combination with negative epistasis between two loci under sexually antagonistic selection is strong enough to drive sex differences in recombination rates due to a reduced recombination in the sex with the strongest value of epistasis (heterochiasmy) (Lenormand 2003). In fact, heterochiasmy is predicted to evolve even in the absence of epistasis if haploid selection on alleles differs between males and females and if the two loci are in linkage disequilibrium because of some other mechanism. In addition, sex-specific selection on haploid gametes may also predict reduced recombination rates on sex chromosomes and an enrichment for haploid-expressed genes on the sex chromosomes (Scott & Otto 2017). Furthermore, haploid selection can drive transitions between male and female heterogametic sex determination systems even if the linkage between the sex-determining locus and the sexually antagonistic locus is not tightly linked, a strict requirement for such transition with loci under purely diploid selection (M.F. Scott et al. 2018). Such transitions may affect population sex ratios, which may increase or decrease with the spread of a new sex chromosome, suggesting that new sex chromosomes may evolve without selection for balancing sex ratio. In addition, purifying selection on male haploid gametes may explain stable sex ratio distortion as a result of the balance between the advantages of purging deleterious mutations via haploid selection and the disadvantages of haploid selection on the sex ratio (Hough et al. 2013).

**GAMETIC HAPLOID SELECTION: EMPIRICAL EVIDENCE**

Although the potential importance of haploid selection for evolutionary—and, more broadly, biological—processes is indisputable, as shown by the theoretical work described in the previous
section, empirical studies investigating haploid selection are still scarce. At this stage, it is important to distinguish between male and female gametes. Already, Haldane (1924) had noted that the scope for selection in female gametes may be much smaller than that in male gametes, because a majority of mature eggs/ovules produced by a female stand a high chance to be fertilized, whereas a tiny fraction of male gametes will fertilize an egg/ovule. I therefore discuss the potential for haploid selection in the two sexes separately. In addition, our understanding of haploid gametic selection differs significantly between plants and animals. Haploid gametic selection is widely accepted in plants—in fact, so much so that it has found its applications in agricultural practices and crop breeding (e.g., Zamir & Vallejos 1983, Clarke et al. 2004, Dominguez et al. 2005). The reason for this is that mitotic cell division and growth are necessary during pollen tube growth, leading to the expression of up to 60% of genes in haploid pollen (Haldane 1932, Mascarenhas 1990, Wallbot & Evans 2003, Borg et al. 2009), of which approximately 10% are expressed exclusively in the haploid stage (Honys & Twell 2004, Borg et al. 2009, Arunkumar et al. 2013).

In contrast, a persistent view holds that the scope for haploid selection in animals is minimal (but see Joseph & Kirkpatrick 2004). The sperm phenotype and its ability to fertilize an egg determine the fitness of the diploid male, and in particular, males facing sperm competition are selected to produce competitive sperm phenotypes (Birkhead & Møller 1998). As mentioned above, due to the potential conflict arising from differential selection in haploid gametes and diploid organisms, a general assumption is that the diploid male has control over sperm phenotypes and that the haploid sperm genotype has a minimal impact on the sperm phenotype. This view was confirmed early on by empirical findings that Drosophila and mice mutant males producing sperm that lacked a nucleus were still capable of fertilizing an egg (Muller & Settles 1927, Lindsley & Grell 1969, Lyon et al. 1972). In addition, even those genes expressed in early haploid spermatids appeared to be shared through cytoplasmic bridges, rendering the spermatids functionally diploid (e.g., Dym & Fawcett 1971, Caldwell & Handel 1991). A statement in the foreword to the Proceedings of the Symposium on Gamete Competition in Plants and Animals in 1975 saying “haploid transcription is apparently a relatively rare and insignificant phenomenon among higher animals” (Mulcahy 1975) epitomized the general idea that little opportunity exists for haploid selection to occur. However, an increasing body of empirical evidence is continuously challenging and overturning this view.

**Haploid Gene Expression**

Active transcription of genes at the haploid postmeiotic stages is a key prerequisite for haploid gametic selection to occur. As mentioned above, haploid gene expression in plants, and hence selection during pollen tube growth, is extensive (e.g., Mascarenhas 1990, Sari-Gorla & Frova 1997). In animals, evidence for postmeiotic gene expression is more complicated to obtain; however, more recent studies of postmeiotic transcription have since revised the prevailing view, and postmeiotic gene expression has been reported in several taxa including Drosophila and a range of mammals (reviewed in Erickson 1990, Hecht 1998, Steger 1999, Kanippayoor et al. 2013).

A possible explanation for the fertilization ability of genome-depleted sperm in earlier experiments is the syncytial organization of early spermatids, in which several (more than 100 in mammals and 64 in Drosophila melanogaster) haploid cells stay connected by cytoplasmic bridges, allowing the distribution of transcripts (Erickson 1973). This characteristic suggests that the sharing of nascent RNAs among early spermatids renders these cells functionally diploid. Although evidence for sharing is strong for some genes, such as the allelic-expressed protamines hProt1 and hProt2 in humans and similar alleles in mice, the sharing is not always perfect, and expression biases may persist (Kanippayoor et al. 2013). The lack of sharing and the compartmentalization and immediate translation of mRNA of the house mouse gene spam1, for example, supports the idea of haploid expression of this gene associated with male infertility (Zheng et al. 2001).
Previous estimates suggested that up to several hundred genes are expressed at the postmeiotic stage in several taxa (Joseph & Kirkpatrick 2004). An important next step is to take a broader approach and create transcriptome maps of nascent RNAs in postmeiotic spermatid cells to identify all allele-expressed genes.

The transcriptional and translational activity in mature sperm is another much debated topic (Ren et al. 2017). The first evidence for active translation in mature sperm came from bovine sperm that were observed to incorporate radio-labeled ribonucleoside triphosphates into RNA and protein molecules (Premkumar & Bhargava 1972, MacLaughlin & Terner 1973). The general thought, however, was that this translational activity was confined to the sperm mitochondria (Miller & Ostermeier 2006), until a more recent study demonstrated translational activity in mature sperm of several mammalian species and showed that mammalian sperm translate nuclear-encoded proteins by mitochondrial-type ribosomes (Gur & Breitbart 2006, 2008). Such translational activity in late spermiogenesis and in mature sperm might also explain the striking finding of eight proteins differing in abundance in bull sperm bearing both X and Y chromosomes (C. Scott et al. 2018).

Patterns of transcriptional and translational profiles in the haploid postmeiotic phases of spermiogenesis are likely to vary markedly across species. Spermatogenesis exhibits remarkable variation across taxonomic groups in the way germ cells are organized and divide before, during, and after meiosis; in the genes that are involved; and most likely also in the transcriptional and translation activity in postmeiotic spermatids and mature sperm (Ramm et al. 2014). Key characteristics with potential relevance to our understanding of haploid selection that varies across species are the way the sperm nucleus is condensed and how histones are replaced during sperm maturation: Histones are retained to varying degrees ranging from an estimated 1% in human sperm, 10–50% in house mouse sperm, and up to 100% in zebrafish sperm. The specific location and nature of different histone variants may provide further information about the functional activity of sperm nuclei and how it may vary across species.

Furthermore, the condensation of the sperm nucleus occurs in a highly organized manner, exhibiting a distinct hairpin structure with DNA loops similar to those in the chromosome structure found in condensed mitotic cells (Ward 2018). Such conserved organization results in specific genome regions consistently being located on the surface of the nucleus, whereas other regions are hidden within the nucleus and may be inaccessible to any transcription and/or translation factors. Whether the specific location within the nucleus has any significance for the activity of genes is currently speculative at best but certainly deserves further attention.

Genetic Diversity and Genotype–Phenotype Links Among Male Gametes

The processes of segregation and recombination occurring during most meiotic processes of spermatogenesis lead to the prediction of substantial genetic variation among sperm within an ejaculate. The number of chromosomes determines the variation among sperm due to segregation, and the rate of male recombination adds an additional level of variation. The latter may be of particular importance if genes under haploid selection show any additive effects or signs of epistasis. Cohen (1967, 1973) reported a positive association between the recombination rate (chiasma rate) and the number of sperm produced across a range of species. Cohen’s interpretation of this finding was that recombination events may lead to the production of suboptimal sperm, and to compensate for the number of suboptimal sperm, males produce relatively more sperm. Data on recombination events are now available at a higher resolution, and his observation can be tested on a larger sample size. More generally, single-sperm genotyping may be a great tool to assess genetic variation among sperm within a male’s ejaculate and get a better estimate of the recombination rate and its...
role in generating genetic variation among sperm (e.g., Lu et al. 2012, Wang et al. 2012, Xu et al. 2015) (see also the section titled Genomic Signatures of Haploid Gametic Selection).

A logical next question is then: What is the evidence for the role of this genetic variation to contribute to the phenotypic diversity among sperm within an ejaculate? A study in the zebrafish reported allelic divergence across the entire genome between sperm pools selected for different longevity phenotypes from the ejaculate of individual males (Alavioon et al. 2017). Similarly, in the Astyanax cave fish, sperm carrying specific alleles exhibited a different reaction phenotype in response to exposure to Hoechst dye (Borowsky et al. 2018). Moreover, X- and Y-carrying bull sperm differ in eight mitochondrial membrane proteins, ranging from cytoskeleton protein–coding genes to nicotinamide adenine dinucleotide (NADH) dehydrogenases involved in adenosine triphosphate (ATP) production via oxidative phosphorylation and acetylcoenzyme A carboxylase, both of which play key roles in energy production (C. Scott et al. 2018). All these very recent findings support the idea that the haploid genome does play a role in determining sperm phenotypes, but identifying which exact genes are involved needs further investigation. A first tentative conclusion we may draw here is that phenotypic traits affected by the haploid genome generally seem to be functional phenotypes rather than morphological structures.

The genetic diversity among sperm is certainly not the only mechanism affecting phenotypic variation among the sperm of one male. Even in the mitotically produced, genetically identical sperm of haploid males in eusocial insects, morphological variation is substantial, suggesting that the phenotypic variation in size is largely determined by variation during spermatogenesis rather than being affected by the sperm haplotype (Fitzpatrick & Baer 2011). Similarly, an elegant experiment performing crosses between Drosophila lines selected for long and short sperm suggested that the haploid sperm genome has little influence on the length of sperm, as no distinct size classes were found among the male F1 crosses (Pitnick et al. 2009). These results further support the idea that at least morphological sperm traits are likely to be under the control of the diploid male and reflect variation in male condition (Holt & Van Look 2004).

Evidence for Impact of Haploid Gametic Selection on Offspring Phenotypes

Gametes from one male are expected to compete (Manning & Chamberlain 1994, Haig & Bergstrom 1995), and some gametes may have an advantage at fertilizing eggs over their sibling gametes. In plants, competition among pollen of different males and also within the same male have been shown to not only affect fertilization success rates but also directly improve offspring fitness (e.g., Snow & Spira 1996, Aronen et al. 2002, Lankinen et al. 2009). In addition, within-plant pollen competition appears to reduce inbreeding depression in Dalechampia sandens (Armbruster & Gobeille Rogers 2004). Finally, selection on pollen tube growth occurring during fertilization has enabled crop breeders to obtain cold-tolerant chickpea Cicer arietinum and tomato Solanum lycopersicum plants in the next generation (Clarke et al. 2004, Domínguez et al. 2005). In contrast, evidence for competition and selection among haploid sperm affecting offspring fitness is less abundant.

Evidence for a role of selection on sperm genotypes has been reported in the house mouse Mus musculus, for which sperm reaching the fallopian tube showed lower levels of DNA fragmentation (Hourcade et al. 2010), and in the boar Sus scrofa domesticus, for which chromatin-unstable sperm were less likely to reach oocytes in vivo (Ardon et al. 2008). These two examples suggest that selection on the ability of sperm to reach the site of fertilization is at least partly associated with the genetic quality of the sperm and prevents suboptimal sperm from fertilizing an egg. The most direct evidence for a link between the sperm phenotype and the sperm genotype comes from two
recent studies in fish. In the Mexican cave fish *Astyanax* spp., sperm haplotypes seem to be directly associated with phenotypic variance, in which sperm from a hybrid male could be distinguished into two phenotypic groups (Borowsky et al. 2018). Similarly, a study in the zebrafish *Danio rerio* showed not only a direct link between sperm phenotypes and offspring fitness from early life into adulthood but also a link between the haploid sperm genotype and the selected sperm phenotype (Alavioon et al. 2017). Signs of allelic divergence between two sperm pools selected for differential phenotypes from within the same ejaculate were found across the entire genome, although some of these signals may be false positives due to the effects of linkage and background selection. A better understanding of the expected genetic variation and the role of linkage in determining genetic variation among sperm (and gametes more generally) is needed to improve the accuracy of such analyses.

**Haploid Selection in Female Gametes**

As mentioned above, the opportunity among female gametes generally is thought to be much more limited for three reasons—namely, the costs of female gamete production, the high fertilization success in female gametes, and the delayed completion of the second meiotic division shortly before, during, or after fertilization (Immler & Otto 2018). Nevertheless, there are several instances for which selection among female gametes may still be of importance, and they are therefore worth further consideration.

A possible opportunity for haploid selection on female gametes occurs precisely during fertilization and the completion of the second meiotic division. Arresting the second meiotic division in the metaphase until a sperm enters the egg provides an opportunity for the female to assess the genetic match between the paternal and the maternal genomes forming the diploid zygote. Direct empirical studies of such processes are currently lacking, but an interaction between male and female haploid pronuclei has been suggested in the context of several findings. The perhaps most striking example of a possible process of choice at the haploid level comes from a marine invertebrate, the comb jelly *Beroe ovata*. In this species, polyspermy leads to the presence of several sperm pronuclei in one egg, and the female pronucleus has been observed to migrate among these (Carré & Sardet 1984) (for a video clip of the striking process, see [https://jellybiologist.com/2013/05/20/video-can-an-egg-choose-the-sperm-it-likes-best/](https://jellybiologist.com/2013/05/20/video-can-an-egg-choose-the-sperm-it-likes-best/)). Physiological polyspermy—as opposed to pathological polyspermy, which is detrimental to the development of the zygote—is widespread across taxa and in many cases appears to be obligatory for successful fertilization and development of an egg (see Snook et al. 2011 for review). However, the reason for physiological polyspermy is unclear, and the scope for a possible mate choice process occurring at this stage warrants further careful investigation.

Direct empirical tests for possible evidence of assortative fusion based on the haploid gametes are scarce and come exclusively from animals. Two studies in whitefish *Coregonus* spp. and Atlantic salmon found no evidence of assortative fusion with respect to different major histocompatibility complex (MHC) alleles (Wedekind et al. 2004, Promerová et al. 2017), whereas a study in the three-spined stickleback *Gasterosteus aculeatus* reported a possible role of assortative fusion to optimize MHC genotypes in resulting offspring (Lenz et al. 2018). Similarly, in the house mouse *M. musculus*, a possible interaction between the male and the female haploid genomes has been suggested to explain non-Mendelian inheritance of certain alleles (Wedekind et al. 1996, Nadeau 2017). Nadeau (2017) suggested that these haploid genome interactions could be condition dependent. Such condition dependence more generally would further explain the difficulty to provide firm and clear evidence for haploid selection to occur. However, at this stage, this idea is still rather speculative and needs careful further testing.
Evidence for Roles of Haploid Selection in Evolutionary Processes

Gametic haploid selection has been suggested to play a potentially important role in a range of evolutionary processes including adaptation, purifying selection, the evolution of sex chromosomes, and heterochiasmy (see the section titled Haploid Selection in Theory) and is increasingly mentioned in empirical studies as a possible underlying mechanism. For example, a lack of sex bias in expression of mitonuclear genes across several species was thought to be a possible consequence of gametic haploid selection due to the potentially tight interaction between the sperm and its mitochondria (Dean et al. 2014). The reasoning was that individual sperm can only carry either an X or a Y chromosome, and haploid expression of mitonuclear genes would select against sex linkage, as such linkage would render 50% of sperm nonfunctional. More recently, the role of gametic haploid selection in the evolution of sex chromosomes was demonstrated in two Rumex species (Sandler et al. 2018). By examining the evolution of gene expression in flower buds and pollen to test for signatures of haploid selection acting during plant sex chromosome evolution, the authors found a bias of genes with high ancestral pollen expression bias on sex chromosomes compared with autosomes. Furthermore, genes on the Y chromosome were more likely to become enriched for pollen-expressed genes with a significant loss of genes with low pollen expression levels. Haploid selection was also thought to be the most likely explanation for the observed variation of heterochiasmy across a range of plant species (Lenormand & Dutheil 2005). Similarly, it seemed the most plausible explanation for observed heterochiasmy in a coral Acropora millepora and may also explain a low polymorphism level of one linkage group in the male (Wang et al. 2009). In this study, the male genotype was inferred from its sperm sample, and haploid selection may result in a difference at some loci between the ejaculate genotype and the genotype in the adult male. This difference may occur, for example, if one homologous chromosome corresponding to one linkage group produces functional sperm, whereas the other chromosome carries deleterious alleles resulting in nonfunctional sperm. These ideas are untested at this moment, but they deserve further attention.

All the evidence for gametic haploid selection discussed in the sections above is focused on the possible mechanisms allowing for genes to be expressed in haploid postmeiotic cells, evidence that haploid genotypes affect pollen and sperm phenotypes and how such phenotypic variation may affect offspring phenotypes. An important additional and more recent approach to look for signs of haploid selection is to take advantage of novel sequencing technologies to estimate the strength and nature of selection occurring across the genome (see the section titled Genomic Signatures of Haploid Gametic Selection). By using a combination of quantitative genetics, population genetics, and genomics, assessing the true potential for haploid selection in animal male gametes will be more feasible than ever before.

Genomic Signatures of Haploid Gametic Selection

With the advent of continuously improving sequencing technologies, we may take entirely new approaches to assess and understand the importance of selection on haploid genomes in diploid organisms. Whole-genome sequencing is a powerful tool to study rates of (a) mutation, (b) recombination, and (c) selection—all key ingredients for understanding population ecology and evolution.

Long-read sequencing technologies allow a more accurate identification of true de novo mutations and improve our estimates of mutation rates. Another promising step is to move to a single-cell level, and although this is currently not quite possible, assessing mutation rates at the level of gametes and comparing them with mutation rates in pedigrees will provide estimates of purifying selection.
Similarly, single-cell sequencing improves our estimates of recombination rates and our understanding of the genetic variation present among sperm/pollen due to segregation and recombination. Recombination estimates from pedigrees are likely to underestimate the frequency of recombination events, particularly if these events have deleterious effects due to disrupted allele combinations.

Sequencing technologies also enable us to identify signs of selection. Pioneering studies in the outcrossing plant species *Capsella grandiflora* and *Arabidopsis thaliana* have provided invaluable insights into what we may expect to find in genes exclusively expressed during the haploid phase, for genes expressed in both phases, and in genes expressed purely in the diploid phase (Arunkumar et al. 2013, Gossmann et al. 2014). In both studies, haploid-exclusive genes had more sites under strong purifying selection, a greater proportion of adaptive substitutions, and faster protein evolution compared with genes exclusively expressed in the diploid phase. The effect of purifying selection against strongly deleterious mutations also persisted in genes expressed during both phases, whereas the signs of directional selection were only marginally higher in genes with biphasic expression compared with diploid-exclusive genes.

Based on the recent technical advances and the findings in plants, it may be worth looking at rates of mutation and selection in animals and comparing genes expressed in haploid spermatids with those expressed exclusively in the diploid organism. In fact, a study on human populations reported a surprisingly high signal of purifying selection in such spermatid-expressed genes, and the authors stated: “Although there is no guarantee that gametic selection is beneficial to the organism, if significant purging can occur during gametogenesis (even at the diploid germ-cell stage), selection at this phase can dramatically enrich the proportion of ‘purified’ genomes for fertilization” (Reed & Aquadro 2006, p. 481). Furthermore, genes involved in the condensation of spermaticogenic DNA expressed in haploid spermatids are often rapidly evolving between mammalian species (Good & Nachman 2005), and more generally, rates of protein evolution have been found to be positively correlated with developmental timing of expression of genes involved in spermatogenesis (Podlaha & Zhang 2003, Good & Nachman 2005, Podlaha et al. 2005). A striking example of extremely high levels of insertion–deletion variation of an alanine-rich repetitive motif in natural populations of *Mus musculus domesticus* and *Mus musculus musculus* was found at Testis-specific gene a8 (*Tsga8*), a spermatogenesis-specific gene expressed during postmeiotic chromatin condensation and nuclear transformation (Good et al. 2011).

Whether haploid expression is also (partly) responsible for the rapid evolution in reproductive genes more generally (Swanson & Vacquier 2002) needs further investigation. The fact that the ratio of sites with nonsynonymous versus synonymous substitutions (dN/dS) exceeds 1 for genes whose products are found in mature sperm (Ezawa & Innan 2013), combined with the positive correlation between high dN/dS value expression during the postmeiotic haploid phases (Good & Nachman 2005) and the X/Y-sperm specific protein phenotypes (C. Scott et al. 2018), may indicate a possible role for haploid selection. Combining sequencing technologies with targeted experimental crossing and experimental evolution is likely to provide exciting new insights.

### OTHER HAPLOID GENOMES IN “DIPLOID” ORGANISMS

Besides the haploid gametic phase in sexually reproducing eukaryotes, diploid organisms carry a range of other haploid genomes and genes. Although by definition “diploid” organisms are carrying two sets of alleles in most cells and biallelic gene expression is generally the norm, haploid allelic expression of genes can be found at many different levels. Monoallelic expression may be found in imprinted genes, genes located on sex chromosomes and on chromosomes in haplodiploid species. Furthermore, eukaryotic cells rely on the haploid genomes of organelles such as...
Table 1  Summary of empirical evidence for haploid gene expression and possible consequences such as elevated levels of purifying selection, positive selection, and direct fitness effects of such haploid selection

<table>
<thead>
<tr>
<th>Genome</th>
<th>Haploid expression</th>
<th>Purifying selection</th>
<th>Positive selection</th>
<th>Fitness effects$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gamete$^a$</td>
<td>Yes/Yes</td>
<td>Yes/Unknown</td>
<td>Yes/Yes?</td>
<td>Yes/Yes?</td>
</tr>
<tr>
<td>Female gamete</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear/Yes?</td>
</tr>
<tr>
<td>Paired sex chromosome with inactivation (X)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes?</td>
</tr>
<tr>
<td>Paired sex chromosome in heterogametic sex (X, Z)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Unpaired sex chromosome (Y, W)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Randomly imprinted genomic regions</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Organelle genomes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
</tr>
</tbody>
</table>

$^a$For male gametes, the first indication is for evidence in plants, and the second for evidence in animals. Evolutionary rates of male-specific genes in haplodiploid organisms are comparable with those for unpaired sex chromosomes and paired chromosomes in the heterogametic sex in the absence of inactivation. In contrast, genes under random imprinting on autosomes and sex chromosomes do not show an increase in purifying selection.

$^b$The “Fitness effects” column refers to positive effects on organismal fitness as a result of purifying and/or positive selection on the haploid genome. Evidence in mitochondria and plastids focuses on signs of selection at the mitochondrial level (i.e., among individual mitochondria) rather than at the organismal level.

Haploidy Through Imprinting

Genetic imprinting and gene silencing may render individual loci or entire genomic regions or chromosomes effectively haploid. Random monoallelic expression exists at the genome-wide scale and affects a large number of coding genes (Chess 2012). Random monoallelic imprinting is expected to slow down purifying selection, compared with alleles with complete dominance, but may increase the rate of positive selection, compared with loci with biallelic expression. Sex-specific imprinting, on the other hand, consistently silences loci depending on their parental origin (Reik & Walter 2001). Sex-specific imprinting has been extensively discussed, particularly in the context of sexual conflict over gene expression between the paternal and the maternal genome in zygotes (e.g., Haig 2000). The gene MEDEA is essential for seed development and is a maternally imprinted gene with a silenced paternal allele resulting in haploid expression. As expected for haploid-expressed alleles, MEDEA shows signs of accelerated selection in the outcrossing species Arabidopsis lyrata (Spillane et al. 2007). Sex-specific expression and random monoallelic expression are likely to show similar patterns of purifying and positive selection, if sex-specific imprinting varies across individuals.

Imprinting is also a mechanism for dosage compensation to account for the ploidy differences in the sex chromosomes (X or Z) between the homogametic and heterogametic sexes. In mammals, for example, dosage compensation is achieved by the random inactivation of one of the two X chromosomes in females through DNA methylation, histone modification, and large-scale chromatin restructuring (Brockdorff & Turner 2015). Similar to the observations made for randomly
silenced loci described above, purifying selection in inactivated genes on the X chromosome is reduced, compared with X-linked genes that escape inactivation (Park et al. 2010). The reason for this reduced level of purifying selection is that random inactivation allows deleterious alleles to hide from selection. Interestingly, however, positive selection does not seem to differ between the inactivated loci and loci escaping inactivation in X-linked genes.

**Haploid Chromosomes and Genomes**

Heterogametic sex determination renders one of the two sex chromosomes consistently haploid (Y and W) and the other one (X and Z) haploid one-third of the time in the heterogametic sex. Concentrating entirely on the ploidy state of each sex chromosome, we would expect purifying selection to be enhanced on Y and W compared with X and Z chromosomes. However, these predictions are strongly confounded by the fact that heterogametic sex chromosomes evolve asexually; hence, recombination as an efficient means to remove deleterious mutations from chromosomes does not occur. Furthermore, factors such as reduced effective population size (Charlesworth 1978, Gordo & Charlesworth 2000) and background selection (Charlesworth 1996) should further contribute to the degeneration of the heterogametic sex chromosome and make the generation of predictions for the Y chromosome evolution challenging. Nevertheless, genome sequencing data from 11 genes in 9 *Drosophila* species support the idea of recurrent events of positive selection and uncompromised purifying selection against strongly deleterious mutations (Singh et al. 2014). Only mildly deleterious mutations appear to be maintained by background selection.

X and Z chromosomes exist in a haploid state in one-third of their evolutionary time either in company of a heterogametic nonrecombining Y or W or without any homologous chromosome in species with X0 sex determination systems. Genes expressed during this time are again predicted to experience enhanced levels of purifying and positive selection compared with diploid autosomes. These predictions are confirmed in the sequencing data from 12 *Drosophila* species, for which elevated biased codon usage in genes on the X chromosomes supports the hypothesis of increased purifying selection, whereas increased substitution rates as a sign of positive selection were observed only in some lineages (Singh et al. 2008).

The selection dynamics described above for sex chromosomes apply also to species with haplodiploid sex determination (haploid males and diploid females). However, any genes exclusively expressed in males will experience only haploid selection, and genes expressed exclusively in females only diploid selection. The low effective population sizes characteristic for eusocial haplodiploid insects may be a confounding factor when estimating strength of selection (Romiguier et al. 2014). Nevertheless, haploid males in eusocial insects may be the equivalent of the haploid gametes produced by a diploid male. Studying the genomic signatures of selection in eusocial insects may therefore reveal important information about what we would expect in one standard fertilization event between a male and a female.

**Haploid Organelles**

The haploid genomes of organelles typically found in eukaryote organisms provide an additional opportunity to understand the genomic evolution in response to haploid selection. Organelle genes as those found in mitochondria and plastids have been regularly used for phylogenetic reconstructions across taxa under the erroneous assumption that these genes may serve as neutral markers due to a lack of selection. However, this view has since been changed, and the fundamental differences between nuclear and organelle genes have been increasingly recognized (Ballard & Whitlock 2004, de Vries & Archibald 2018). The evolution of mitochondria and plastids may be
assessed at two levels, one at the individual organelle level and one at the organismal level of the individual carrying the organelle in its eukaryotic cells. The seemingly rapid evolution of mitochondrial genes, for example, may be explained partially by the nature of its haploid genome, on the one hand, and partially by the vast population sizes, on the other hand. Each eukaryotic cell may contain up to 2,000 mitochondria, and these numbers duplicate during each mitotic division. Such vast numbers are expected to show rapid rates of evolution even in asexually reproducing populations, particularly given that positive selection and purifying selection are expected to be strong in these haploid genomes.

However, estimates of mitochondrial evolution are generally performed at the organismal level. This approach may not be a due reflection of the processes affecting the evolution of mitochondria (and other organelle) genomes. Recombination rates in mitochondria vary markedly and may range from nonrecombinant in mammals to relatively high rates of recombination in some species of yeast and in species with biparental mitochondrial inheritance (Barr et al. 2005). Mitochondria within an individual are rarely homogeneous but in effect exhibit substantial genetic variation and bottlenecks, as those assumed to occur during inheritance are likely to be nowhere as severe as assumed and as a single oocyte may carry over as many as $10^6$ mitochondria (Shoubridge & Wai 2007, Radzvilavicius et al. 2016). In addition, (deleterious) mitochondrial mutations have only a noticeable effect on organismal fitness at frequencies as high as 60–80% (Rossignol et al. 2003). Evolutionary constraints affecting mitochondrial evolution may be caused by the tight mitochondrial interactions. Estimates of purifying selection and directional selection as those reported, for example, in mammals may therefore not directly reflect the population genetic processes occurring at the mitochondrial level. However, they will still be important for understanding large-scale evolutionary processes. Given the complexity of organelle biology, more in-depth investigations are required to fully understand the genetic and genomic dynamics explaining the evolutionary pathways described across species at the macroevolutionary level.

CONCLUSIONS AND FUTURE DIRECTIONS

Although absolute values of strength of selection and rates of evolution vary across the studies investigating the role of haploid expression for purifying and positive selection, the qualitative patterns emerging from across the different haploid genomes (Table 1) are surprisingly similar. The role of haploid selection for purifying and positive selection finds strong support in many of the scenarios discussed above, including haploid gametic selection, imprinted genes, and haploid selection on sex chromosomes. The complex biology and the multiple levels of selection experienced by organelles render the qualitative and quantitative description of evolutionary patterns more challenging. However, research is heading in exciting directions, and intriguing new data should soon be available to answer the questions raised in the previous sections.

This review clearly demonstrates that the effects of haploid selection at all levels may have striking effects for evolutionary processes in diploid organisms. In particular, haploid gametic selection may be more important in animals (and plants) than assumed so far. To obtain a complete picture, questions about the mechanisms maintaining genetic variation at loci under haploid selection are among the most important to be addressed. One possibility is that antagonistic selection across the ploidy levels and/or the sexes maintains a stable polymorphism. Balancing selection is a non-mutually exclusive alternative mechanism maintaining genetic variation at loci under haploid selection. Scenarios of balancing selection may be found in changing conditions during fertilization events affecting gene expression and metabolic rates in both pollen and sperm. These scenarios lead to the next question about the nature of the genes expressed during the haploid phase. In plants, these are generally housekeeping genes (Arunkumar et al. 2013), whereas in animals such
information is currently missing. Additive genetic variation and epistasis among such genes may result in distinct sperm cohorts, for which more than one allele combination can be optimal. In addition, soft selection among gametes produced by one individual may be an additional factor that helps maintain genetic variation in a population. However, these ideas are currently untested in both plants and animals.

Overall, understanding the evolutionary dynamics in the different haploid genomes described here will substantially contribute to our understanding of evolutionary processes as a whole. The identification of the importance of purifying and directional selection at the haploid gametic stages, for example, may help provide the information needed to link apparent discrepancies for mutation rate estimates based on phylogenetic data sets compared with those at the individual and family level. It may also provide explanations for the maintenance of genetic variation despite apparently strong selection on the diploid organisms and may explain why the signs of inbreeding depression are often not as severe as expected even in small populations of endangered organisms.

**DISCLOSURE STATEMENT**

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

**ACKNOWLEDGMENTS**

I would like to thank Mike Whitlock for giving me the opportunity to write this review and for his invaluable feedback. I am particularly grateful to Glyn McAlister and the Norfolk Fire and Rescue Service for preventing our house from burning down while writing this article. S.I. is funded by grants from the European Research Council (HapSelA-336633), the Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellowship), and the Human Frontier Science Program.

**LITERATURE CITED**


Ezawa K, Innan H. 2013. Competition between the sperm of a single male can increase the evolutionary rate of haploid expressed genes. *Genetico* 194(3):709–19


Kanippayoor RL, Alpern JHM, Moehring AJ. 2013. Protamines and spermatogenesis in *Drosophila* and *Homo sapiens.* *Spermatogenesis* 3(2):e24376


